GENERAL ETHICS

Formal and effective autonomy in healthcare

A P Schwab

J Med Ethics 2006;**32**:575–579. doi: 10.1136/jme.2005.013391

This essay lays the groundwork for a novel conception of autonomy that may be called ''effective autonomy''—a conception designed to be genuinely action guiding in bioethics. As empirical psychology research on the heuristics and biases approach shows, decision making commonly fails to correspond to people's desires because of the biases arising from bounded cognition. People who are classified as autonomous on contemporary philosophical accounts may fail to be effectively autonomous because their decisions are uncoupled from their autonomous desires. Accordingly, continuing attempts to value patient autonomy must go beyond existing philosophical conceptions of autonomy to consider the background conditions of human decision making.

Obstacles to good decision making by patients are a serious concern in healthcare. Certain obstacles are generally expected to arise: a lack of information, mental disability, dementia or immaturity. But obstacles also arise from unexpected sources. The following example comes from a study on hypothetical preferences for medical treatments.[1] In this study, respondents were given information on the outcomes of two treatments for lung cancer. Although the statistics presented were identical, they were framed in opposing ways—in terms of survival rates for some respondents and in terms of mortality for others (the following is presented exactly as it was to participants in the study).

> Survival framing:
> Surgery: Of 100 people having surgery 90 live through the post-operative period, 68 are alive at the end of the first year and 34 are alive at the end of five years.
>
> Radiation therapy: Of 100 people having radiation therapy all live through the treatment, 77 are alive at the end of one year and 22 are alive at the end of five years.
> Mortality framing:
>
> Surgery: Of 100 people having surgery 10 die during surgery or the post-operative period, 32 die by the end of the first year and 66 die by the end of five years.
> Radiation therapy: Of 100 people having radiation therapy, none die during treatment, 23 die by the end of one year and 78 die by the end of five years.

According to the principle of invariance in rational choice theory, the rationality of a choice is not affected by the framing of the information—the same statistics presented in different ways should result in the same decision. In this case, however, the number of respondents who favoured radiotherapy went from 18% for those presented with the survival framing to 44% for those presented with the mortality framing.

In this paper, I identify three implications of this study and other research similar to this one for the practice of autonomy in healthcare. Firstly, I show how the biases of human decision making have a major effect on autonomous decision making. Although the desires or interests of a particular person may be autonomous, the heuristics and biases of human decision making may undermine attempts to achieve those desires or interests. Secondly, this leads me to show how existing procedural and substantive accounts of autonomy (which I call formal) do not consider these biases, and I suggest a replacement or modified account that I call ''effective autonomy''. My account of effective autonomy, which explicitly deals with the effect of cognitive biases in human decision making, most resembles the formal accounts of autonomy dubbed as ''procedural'', although I briefly question the value and validity of distinguishing between procedural and substantive accounts. Finally, I offer speculative policy changes for healthcare that follow from effective autonomy. These will be necessarily brief—more research is required to produce anything more than speculation about needed changes.

Effective autonomy is the matching of formally autonomous interests or desires with decisions that will achieve those interests or desires. The term effective autonomy is already prevalent in public discourse, although not in philosophy or bioethics. Its most common use is in political discussions and it is often used to refer to practical autonomy that is not legally recognised. For example, when a group is self-governed despite legal governance by some other body (ie, the Kurds in Northern Iraq), they are said to have effective autonomy. In my use of effective autonomy in healthcare, I am suggesting the opposite situation—we already recognise the legal right to autonomy in the medical encounter, but need to consider the biases of decision making that may undermine effective autonomy.

FORMAL ACCOUNTS OF AUTONOMY

Most regard autonomy as something of value, but many different explanations of its value are defended. Agich[2] notes this in his work *Autonomy and Long-Term Care*, in which he describes both

Correspondence to: Abraham P Schwab, Department of Philosophy, Brooklyn College, 2900 Bedford Ave, Brooklyn, NY 11210, USA; ASchwab@brooklyn.cuny.edu

Received 5 July 2005
In revised form 11 January 2006
Accepted for publication 13 January 2006

the philosophical and cultural perspectives on autonomy. In both cases, he claims that autonomy is ''generally, though not universally, regarded with approval'' (Agich,[2] p 7, 14). In support of Agich's claim, there are, of course, the standard liberal arguments for the value of autonomy and these have been joined (with reservations) by communitarians,[3 4] feminists[5–8] and bioethicists.[9–11] In these existing accounts, the focus has remained on either the formation and identification of desires, beliefs or characteristics or on the whole individual and not on the characteristics of human decision making.[i] I will call these formal accounts of autonomy.

Formal accounts of autonomy share a relatively abstract conceptual focus, including formal requirements that describe the necessary characteristics or capacities of particular people or particular decision making processes. There is a lack of corollary attentiveness to the conditions under which these decisions are made. Such an oversight limits the usefulness of an account for fostering autonomous decisions in healthcare (as well as other areas).

LIMITS OF HUMAN JUDGEMENT

A common concern that motivates many conceptions of autonomy comes from brave-new-world-like scenarios, where people ''believe'' that they are working in their own best interests, but in fact their actions are simply the result of social control. We do not want the decisions of thoroughly brain-washed people to count as autonomous. This demands certain conditions that eliminate obviously poor decisions. Examining how humans make decisions, however, shows that blatant poor decision making can arise much more subtly than through systematic deception. The mistaken assumption is that outside of relatively extreme conditions, people make good decisions and that they are at least instrumentally rational. That is, if a person wants ''P'' and believes that ''not P unless Q'', then that person will try to bring about ''Q.'' Research shows that people do not consistently (or even regularly) make instrumentally rational decisions. The obvious question is why not. In short, the answer is ''bounded cognition''.

Suggested first by Simon, the bounded cognition framework assumes that people attempt to make optimal decisions, but that their decisions are affected by the conditions under which decisions are made.[13 14] These include but are not limited to limits on available information, awareness of relevant criteria, time constraints and limited memory. Importantly, because of the robust conclusions shown in the empirical research in this area, we can say more than that human decisions are biased—they are biased in predictable ways.[15]

One of the key concepts for this theory of human decision making is the idea of a heuristic. A heuristic is a short-cut or rule of thumb in a decision-making process. For example, we operate under the condition of limited time. We do not have the opportunity to gather all the information that could be relevant to a decision we need to make. Instead of (literally) randomly choosing, we use a heuristic. This makes it possible for us to make decisions at a relatively rapid pace, but it also increases the likelihood that we will make mistakes. Heuristics are not bad or inefficient—in many cases they are much more efficient than laboriously researching every decision. None the less, heuristics open up decision making to certain likely mistakes.

At least three general categories of heuristics exist: representativeness, availability, and anchoring and adjustment.[16–18] Each heuristic opens up the possibility of several biases, but to make the case for effective autonomy, I will focus only on the heuristic of representativeness. In short, the representativeness heuristic arises from similarities in occurrences, outcomes and classifications. Although this heuristic is useful in many instances, it is also overapplied in many others, leading to biases in judgement. For example, a ''fair'' coin will, on average, end up with a 50–50 split of heads and tails. People may expect a fair coin to flip heads, then tails, then heads, then tails, etc. In fact, a fair coin will not always (or even often) flip in this way. Misapplying an objective characteristic (eg, a ''fair'' coin) to a small data sample is called the ''insensitivity to sample size'' bias arising from the representativeness heuristic.

This heuristic can also lead to particularly problematic judgements when it is used by a group of professionals. For example, evidence indicates that obese people are at greater risk for certain complications. This evidence has been used by practitioners to justify excluding obese people from surgical interventions. Everett *et al*,[19] however, indicate that this claim is false in at least some cases. Accordingly, deciding against surgery for an obese patient in all cases is still an unjustified judgement. One explanation for the unjustified (but presumed justified) non-treatment of obese people by the medical profession is a biased use of the representativeness heuristic—specifically insensitivity to base rates. Base rates refer to the standard frequency of a given event in a given context. So, for example, there is a base rate for college undergraduates who receive grade A. Because this base rate has changed over the past several decades, a 4.0 grade point average means something different today than it did in the 1960s. As this example shows, failure to consider the effects of bounded cognition can lead to poor decisions and, in healthcare, poor outcomes.

PROCEDURE AND SUBSTANCE: AN UNCERTAIN DISTINCTION

Formal accounts of autonomy have been separated into two camps, procedural and substantive. A procedural account defines autonomy in terms of following certain procedures. If the appropriate procedures are followed, then the outcome is autonomous—regardless of the content of the outcome or the characteristics of the decision maker. Recent examples include procedural requirements such as self-reflection, freedom from inhibitions, identification or endorsement of desires, and at least minimal rationality. Substantive accounts differ from procedural accounts by limiting possible autonomous outcomes. They require certain conclusions or certain perspectives as a condition of autonomous choice. For example, Benson[21] includes the substantive requirement that a person be able to competently criticise alternative courses of action according to the relevant norms. This requirement is substantive because it restricts autonomy to people with certain characteristics. No matter what procedures are followed in achieving the desire, a person unaware of the relevant norms cannot be autonomous.

Substantive accounts of formal autonomy, typified by Benson's account, are prone to devastatingly ambiguous normative standards. Benson[21] initially claimed that for people to take free action, they must have ''an ability to criticize courses of action competently by relevant normative standards''. Benson[22] refined his claim such that autonomous people must view or understand themselves as agents competent to answer for their own conduct. Specifically, these answers would respond to ''normative demands that, from one's point of view, others might appropriately apply to one's actions''.

[i] Berg *et al*[12] provide (or attempt to provide) ''a comprehensive introduction to the theory and practice of informed consent'' for practicing clinicians and, yet, they never mention the psychology of decision making.

Such a substantive account of autonomy is dogged with the indeterminacy of the ''appropriate'' normative standards. Are relevant norms those that an agent views as ''right''? This makes the substantive account so weak as to be useless. If I get to pick and choose what norms to respond to, I will be able to competently respond for all of my actions by simply ignoring those norms that are the most difficult to respond to. Should it be those normative standards that are most widely accepted? It seems not, as this would leave the determination of a person as autonomous and the definition of relevance closely aligned with popularity. The dangers of such a perspective are shown in the familiar history of sexism, racism and classism. Should it be those normative standards that are taken to be true independently of their popularity? Such a standard would require people to know the true from the false (or the relevant from the irrelevant) to qualify as autonomous. Additionally, this requires a method for distinguishing the true from the false beyond currently available standards. And yet Benson never explains how we should determine relevance (or truth). Until such an explanation is offered, Benson's account of autonomy will remain incomplete and therefore impractical.

We have another problem (or another way to get at the same problem) with Benson's view. It is easy to imagine (or remember) people who have defied the norm of ''reason giving'' to justify their actions. Think of the radical who describes the hegemonic views of bourgeois society as requiring reasons that are irrelevant or mistaken. Benson, of course, may defend his view (and accommodate this purportedly autonomous person) by suggesting that some norms are invalid and therefore a person need not be prepared to reply in accordance with them. But, again, this take requires all autonomous people to know which norms are valid and which are invalid.

Christman describes a formal conception of autonomy that I favour, because it avoids the problems of substantive accounts as well as the problems of previous procedural conceptions of autonomy.[23–27] Following the literature, I refer to Christman's account as a procedural one,[22 23 27 28] although, momentarily, I will cast doubt on the distinction between substantive and procedural accounts. Christman identifies four conditions for the procedural autonomy of a desire. Firstly, the formation of a certain desire was not (or would not have been) resisted when the person did reflect (or would have reflected) on the desire. Secondly, this lack of resistance is not due to circumstances that inhibit self-reflection. Thirdly, the self-reflection is minimally rational and does not include self-deception. Finally, at a given time after the autonomous formation of a desire, the person is minimally rational with respect to the desire. Christman's procedural conditions do have some ambiguities. Working definitions of self-deception and the conditions that inhibit reflection should be developed or imported. Moreover, fine-grained standards for manifest inconsistency and resistance (on reflection) would be preferable. These ambiguities, however, are unlikely to be devastating.

I favour something like Christman's procedural account of autonomy over substantive accounts. But Wolf[29] defends a substantive account of autonomy that may be congruous with a procedural account. To be autonomous, Wolf[29] claims that a person must be ''able'' to make decisions in line with ''the True and the Good''. She emphasises that what this entails is a commitment to some degree of ''objective'' evaluative standards. On her definition, even if no definitive conception of the True and the Good is achieved, people, to be autonomous, must be able to govern themselves in accordance with standards independent of their own preferences. This has consistently been characterised as a substantive view of autonomy because it requires an agent to have an objectivist view of normative values. Although no specific values or aims are required, a certain perspective must be embraced for an agent to be considered autonomous.

It is worth asking whether Wolf's substantive view is very different from Christman's procedural view. Wolf requires the ability to govern actions according to the True and the Good, and Christman requires reasoning that is without manifest inconsistency and self-deception. Both require outside judgement for determinations of autonomy. The self-deceived people, by definition, cannot recognise their own limitations and the inconsistency of reasons implicitly acknowledges an objective standard. It certainly seems that Wolf's objectivism will require more than Christman's minimal rationality, but this may result from what we assume must be features of the True and the Good. If we assume that this will include certain political or moral perspectives or certain well-defined commitments, then Wolf's account will go beyond what Christman requires. We may, however, also read Christman's conditions as following from, or even constituting, the True and the Good (eg, a life lived in accordance with the True and the Good is lived without desires and beliefs that have been or would have been resisted, without manifestly inconsistent reasoning and without self-deception). Whether or not these accounts can be happily combined, their potential similarity suggests that the distinction between substantive and procedural accounts of autonomy can be quite blurred.

With this in mind, my perspective on formal autonomy favours Christman's account, but is not committed to a procedural account per se. Regardless, none of the accounts I have discussed (nor any other I can find) deal adequately with the heuristics and biases of human judgement. I find this troubling, as every potentially autonomous decision we encounter is subject to human judgement.

APPLYING BOUNDED COGNITION

As I noted before, bounded cognition marks the limits of formal accounts of autonomy. Even if we suppose that a desire is formally autonomous, because every decision is subject to the biases of human decision making, some or many of people's choices will not reflect their autonomously formed desires. The desire may be formally autonomous, but the decisions may not reflect the desire.

We can, of course, ignore this disjunct and limit ourselves to protecting only formal autonomy. We would commit ourselves to valuing autonomy as a desire formation process and not in the actualisation of the desire. On this view (assuming that it can be adequately defended), achieving autonomously chosen ends would not be as important as merely having (unattained) autonomous desires. Autonomously desiring to have a long life would be important, and actually having a long life would not (or would be less so). Although I cannot deny the value of autonomously desiring a long life, I do think it is better to also have a long life. Tragically, some people will lack the means to have a long life, perhaps because of genetic or economic limitations. Worse, if we do not consider the limits of decision making, these (and other) people will make predictably biased decisions that will further undermine their desire for long life.

Autonomy theorists may agree that effective autonomy is valuable, but argue that effective autonomy is irrelevant to the work that they are doing. In a certain sense, they are right. Defining formal autonomy is important whether or not formally autonomous desires are actualised. But in another sense, this view falls short of appropriate goals of responsible healthcare (and perhaps social policy in general). Although focusing only on formal autonomy will be satisfactory for an abstract theory on autonomy, social practices and policies

require more than formal autonomy. For example, if democracy is preferable on the basis of the best social theories, we need to start by getting a grasp on the formal requirements of a democracy (one of which, coincidentally, is autonomous citizens). At the same time, we miss something important if we value democracy, but remain indifferent to whether or not social conditions (including voting mechanisms or overt coercion) undermine democratic election of public officials. In the same way, if autonomy is valuable, then we should work to define the formal features of autonomy (as so many have done). At the same time, we fail to take patient autonomy seriously and treat it; instead, as an esoteric and intellectual exercise, we remain indifferent to whether or not medical decisions reflect the formally autonomous desires of patients.

Effective autonomy clearly requires formal autonomy, but goes beyond it by requiring that decisions work towards the autonomously identified desires of the decision maker. What is required for an adequate account of autonomy, an account that will foster effective autonomy, is the following condition:

> Effective autonomy is fostered just in case attempts are made to counteract predictable biases resulting from bounded cognition, including, when possible, the appropriate arrangement of background conditions and the use of responsible expertise.

Effective autonomy is fostered when biases resulting from bounded cognition are identified as risks in the decision processes of people and this is followed by ''appropriate debiasing''. What appropriate debiasing means will depend on several variables for each decision, including the likelihood of bias, the effect of bias (eg, towards one alternative or away from another), the seriousness of the decision and the cost of debiasing measures. The approximate value of these variables for the myriad of medical decisions requires empirical research, which has not, to date, been completed.

THE EFFECT OF AFFECT

The role of affect in indicating biased decision making and in debiasing strategies remains unclear. For example, Salovey and Williams-Piehota[30] did not pay methodological attention to the role of affect when they conducted field research on how framing effects influence health-related decisions. This shows that we can identify bias without identifying precisely what role (if any) affect has in particular biases in decision making. Modestly successful debiasing strategies have also been shown in controlled and uncontrolled environments without attention to the role of affect.[31 32]

None the less, growing evidence identifies the specific effects of affect on decision making. Research on affective forecasting (predicting a person's emotional response to a given event) has shown that mistakes in predicting affective responses (from the expected intensity and duration to the expected affect itself) bias decision making in several ways.[33] Lerner and Keltner[34] also showed that specific emotions can have a role in decision making. Finally, Damasio[35] postulates a specific role of affect in the availability bias regarding our views on safety in planes and cars. Much like the effects of framing, the effects of affect (alongside predicted affect) warrant continued study as a set of conditions that can predictably bias decision making. Determining which debiasing strategies will require attention to affect, however, requires a more systematic evaluation of the role of affect in decision biases and corollary debiasing strategies.

In the next section, I suggest speculative policy changes that may arise from incorporating the heuristics and biases approach into a conception of autonomy for medical decision making. My suggestions centre on the bias of overconfidence. Other policy changes will be needed in response to other biases, including those arising from or exacerbated by the affect of decision makers.

SOME SPECULATIVE POLICY CHANGES ON THE BIAS OF OVERCONFIDENCE

It is difficult, and in some cases impossible, to evaluate whether a singular decision outside of a controlled environment is biased. In controlled environments, the biases of a decision can be evaluated because the game can be fixed—we know what the accurate answer is. In the messy areas of healthcare (and many other decisions), the right answer for a single decision cannot always be clearly established in advance or even after the fact. But strategies may be available to help us avoid the predictable pitfalls of decision making. The debiasing strategies to pursue are not meant to foster rationality or do away with irrationality per se, but to produce better outcomes by avoiding the predictable pitfalls in the decision-making process. A previously successful strategy, recommended by Gigerenzer and Hoffrage, required replacing probability formats with frequency formats in describing the likelihood of events (eg, the success of a treatment regimen or the accuracy of a diagnostic test).[36] This has improved the reasoning of naive decision makers through the inclusion of this information in their decision process. In this section, I suggest a couple of strategies that may improve clinical decision making by avoiding the predictable pitfalls of overconfidence.

The bias of overconfidence—the systematic belief that we are right more often than we are—has been shown repeatedly.[37] People believe their judgements to be more accurate than they actually are. In cases where the accuracy of their prediction is important, this can be quite harmful—it may delay when they start chemotherapy, whether they take a drug and whether they consent to surgery. Of course, we cannot a priori decide that patients are mistaken about any particular judgement. Even if we have evidence that this particular patient is extremely overconfident, we remain unable to evaluate any particular judgement. This, however, does not leave us adrift. Hirt and Markman,[38] among others, showed that considering an alternative debiased (to some extent) overconfident judgements about future events.[39] This effect, however, is also limited by the ease of generating alternatives and the need for structure of the individual decision makers.[40]

For medical practice, several possible policy implications arise. Whenever a patient makes a judgement about treatment, a routine prompt to consider multiple alternative outcomes may help in calibrating judgements made by the patient. For example, if a patient is considering chemotherapy and evaluating the side effects, routinely prompting the patient to consider the alternative (the side effects will be more or less difficult or painful or exhausting) may produce better calibrated judgements. Whether or not some medical practitioners already do this, I suggest that this can be an expected routine, or ''standard of care'', for every medical encounter.

Doctors as decision makers are also susceptible to biases of overconfidence. Baumann *et al*[41] provided early evidence of overconfidence among medical practitioners. Again, consider-the-alternative strategies may be successful here, but experts are in a special position to make similar decisions repeatedly. This is important because confidence among experts can be better calibrated by consistent and clear feedback.[42] The need to provide clear and consistent feedback has several policy implications. Requiring an explanation for why another diagnosis is not accurate (and so an implicit

consideration of alternatives) may be a useful routine. Although difficult and expensive, a systems-based policy requiring a review of every patient's outcome by all participating doctors may improve decision making by the doctors.

Although further research may show that these recommendations are not required, inefficient or unsuccessful, we must recognise that a commitment to effective autonomy in medical practice requires attention to the role of predictable and correctable bias in decision making.

CONCLUSION

Effective autonomy is the coupling of decisions with formally autonomous interests—it requires formal autonomy, but goes beyond it by requiring attention to the background conditions in which decision makers will more accurately reflect their desires through their decisions. Effective autonomy is fostered if and only if biases resulting from bounded cognition are identified as risks in the decision-making processes of autonomous people and this is followed by appropriate debiasing.

Competing interests: None.

REFERENCES

1 **McNeil B**, Pauker S, Sox H, *et al*. On the elicitation of preferences for alternative therapies. *N Engl J Med* 1982;**306**:1259–62.
2 **Agich G**. *Autonomy and long-term care*. New York: Oxford University Press, 1993.
3 **Lehman E**, ed. *Autonomy and order*. Lanham, MD: Rowman and Littlefield, 2000.
4 **Etzioni A**. *The new golden rule: community and morality in a democratic society*. New York: Basic Books, 1996.
5 **Friedman M**. *Autonomy, gender, politics*. New York: Oxford University Press, 2003.
6 **Hoagland S**. *Lesbian ethics: toward new value*. Palo Alto, CA: Institute of Lesbian Studies, 1988.
7 **Meyers D**. *Self, society and personal choice*. New York: Columbia University Press, 1989.
8 **Nedelsky J**. Reconceiving autonomy: sources, thoughts, and possibilities. *Yale J Law Feminism* 1989;**1**:7–36.
9 **Schneider C**. *Practice of autonomy: patients, doctors, and medical decisions*. New York: Oxford University Press, 1998.
10 **Wolpe P**. The triumph of autonomy in American bioethics: a sociological view. In: DeVries R, Subedi J, eds. *Bioethics and Society*. Upper Saddle River, NJ: Prentice Hall, 1998.
11 **Katz J**. *The silent world of doctor and patient*. New York: Free Press, 1984.
12 **Berg JW**, Appelbaum P, Lidz CW, Parker. *Informed consent: legal theory and clinical practice*. New York: Oxford University Press 2001, vii.
13 **Simon H**. *Models of man*. New York: Wiley, 1957.
14 **March J**, Simon H. *Organizations*. New York: Wiley, 1958.
15 **Kahneman D**, Slovic P, Tversky A, eds. *Judgement under uncertainty: heuristics and biases*. Cambridge: Cambridge University Press, 2001.
16 **Bazerman M**. *Judgment in managerial decision making*. 5th edn., New York: Wiley 2002.
17 **Nisbett R**, Ross L. *Human inference: strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice Hall, 1980.
18 **Tversky A**, Kahneman D. Judgment under uncertainty: heuristics and biases. *Science* 1974;**185**:1124–31.
19 **Everett E**, Tamimi H, Greer B, *et al*. The effect of body mass index on clinical/pathologic features, surgical morbidity, and outcome in patients with endometrial cancer. *Gynecol Oncol* 2003;**90**:150–7.
20 **Hade A**, Shine A, Kennedy N, *et al*. Both under-nutrition and obesity increase morbidity following liver transplantation. *Irish Med J* 2003;**96**:140–2.
21 **Benson P**. Freedom and value. *J Philos* 1987;**84**:465–86.
22 **Benson P**. Free agency and self worth. *J Philos* 1994;**91**:650–8.
23 **Christman J**. Autonomy and personal history. *Can J Philos* 1991;**21**:1–24.
24 **Frankfurt H**. Freedom of the will and concept of a person. *J Philos* 1971;**68**:829–39.
25 **Thalberg I**. Hierarchical analysis of unfree action. *Can J Philos* 1978;**8**:211–26.
26 **Dworkin G**. *The theory and practice of autonomy*. Cambridge: Cambridge University Press, 1988.
27 **Christman J**. Constructing the inner citadel. *Ethics* 1988;**99**:109–24.
28 **Christman J**. Liberalism and positive freedom. *Ethics* 1991;**101**:343–59.
29 **Wolf SR**. *Freedom within reason*. New York: Oxford University Press, 1990:71.
30 **Salovey P**, Williams-Piehota P. Field experiments in social psychology: message framing and the promotion of health protective behaviors. *Am Behav Scientists* 2004;**47**:488–505.
31 **Renner C**, Renner M. But I thought I knew that: using confidence estimation as a debiasing technique to improve classroom performance. *Appl Cogn Psychol* 2001;**15**:23–32.
32 **Vaughn L**, Weary G. Causal uncertainty and correction of judgments. *J Exp Soc Psychol* 2003;**39**:516–24.
33 **Wilson TD**, Gilbert DT. Affective forecasting. *Adv Exp Soc Psychol* 2003;**35**:345–411.
34 **Lerner JS**, Keltner D. Beyond valence: toward a model of emotion-specific influences on judgement and choice. *Cogn Emot* 2000;**14**:473–93.
35 **Damasio A**. *Descartes' error: emotion, reason, and the human brain*. New York: GP Putman's Sons, 1994.
36 **Gigerenzer G**, Hoffrage U. How to improve Bayesian reasoning without instruction: frequency formats. *Psychol Rev* 1995;**102**:684–704.
37 **Fischhoff B**, Slovic P, Lichtenstein S. Knowing with certainty: the appropriateness of extreme confidence. *J Exp Psychol: Hum Percep Perform* 1977;**3**:552–64.
38 **Hirt ER**, Markman KD. Multiple explanation: a consider-an-alternative strategy for debiasing judgments. *J Personality Soc Psychol* 1995;**69**:1069–86.
39 **Koehler DJ**. Explanation, imagination, and confidence in judgment. *Psychol Bull* 1991;**110**:499–519.
40 **Hirt ER**, Kardes FR, Markman KD. Activating a mental simulation mind-set through generation of alternatives: implications for debiasing in related and unrelated domains. *J Exp Soc Psychol* 2004;**40**:374–83.
41 **Baumann A**, Deber R, Thompson G. Overconfidence among physicians and nurses: the 'micro-certainty, macro-uncertainty' phenomenon *Soc Sci Med* 1991;**32**:167–74.
42 **Arkes HR**, Christensen C, Lai C, *et al*. Two methods of reducing overconfidence. *Organ Behav Hum Dec Proc* 1987;**39**:133–44.